UNDERSTANDING VARIATION IN ESTIMATES OF WEALTH INEQUALITY

ONLINE APPENDIX

MARTA BOCZOŃ

Appendix A. Tables and Figures

Date: April, 2024.

I would like to thank: Jean-Francois Richard, Arie Beresteanu, Roman Liesenfeld, Alistair Wilson, Jesse Bricker, Kevin Moore, William Peterman, John Czajka, Daniel Feenberg, and Allison Schertzer. Remaining errors are my sole responsibility.

TABLE A.1. Regression results of weighted least squares for the wealth shares within the top 10 percent

Regressor	Wealth share of the top k%					
	10%	**5%**	**1%**	**0.5%**	**0.1%**	**0.01%**
	SCF					
Intercept	0.665***	0.546***	0.318***	0.241***	0.114***	0.037***
	(97.556)	(103.860)	(30.716)	(19.407)	(14.223)	(12.592)
Slope	0.067***	0.062***	0.021	0.006	0.004	-0.001
	(6.345)	(7.728)	(1.312)	(0.308)	(0.341)	(-0.123)
R^2	0.890	0.923	0.256	0.019	0.023	0.003
	PUF: Homogeneous					
Intercept	0.648***	0.498***	0.261***	0.195***	0.101***	0.039***
	(168.748)	(122.827)	(69.998)	(57.326)	(37.499)	(19.015)
Slope	0.062***	0.071***	0.075***	0.070***	0.053***	0.030***
	(9.708)	(10.076)	(10.645)	(10.401)	(9.089)	(6.347)
R^2	0.847	0.857	0.870	0.864	0.829	0.703
	PUF: Heterogeneous 10YT					
Intercept	0.647***	0.488***	0.243***	0.172***	0.069***	0.023***
	(112.073)	(71.806)	(32.204)	(29.304)	(22.351)	(14.451)
Slope	0.046***	0.057***	0.065***	0.064***	0.064***	0.034***
	(5.402)	(5.659)	(5.746)	(6.829)	(10.653)	(8.691)
R^2	0.632	0.653	0.660	0.733	0.870	0.816
	PUF: Heterogeneous Moody's					
Intercept	0.653***	0.497***	0.259***	0.187***	0.072***	0.022***
	(99.320)	(62.154)	(28.731)	(28.983)	(21.484)	(13.044)
Slope	0.043***	0.051***	0.055***	0.054***	0.063***	0.035***
	(4.525)	(4.447)	(4.223)	(5.425)	(10.139)	(8.240)
R^2	0.546	0.538	0.512	0.634	0.858	0.800

Note: This table summarizes estimation results from 24 weighted linear regressions of the top-decile wealth shares on a constant and linear time trend. The time variable is defined as 0 for the first observation (1992) and as 1 for the last observation (2010). Consequently, the intercept is the estimate of income share at time zero and the slope is the estimate of the difference in income shares between the first and last observation. I report estimated coefficients and t-statistics in parentheses. "***" denotes statistical significance at the 99 percent significance level, "**" at the 95 percent significance level, and "*" at the 90 percent significance level.

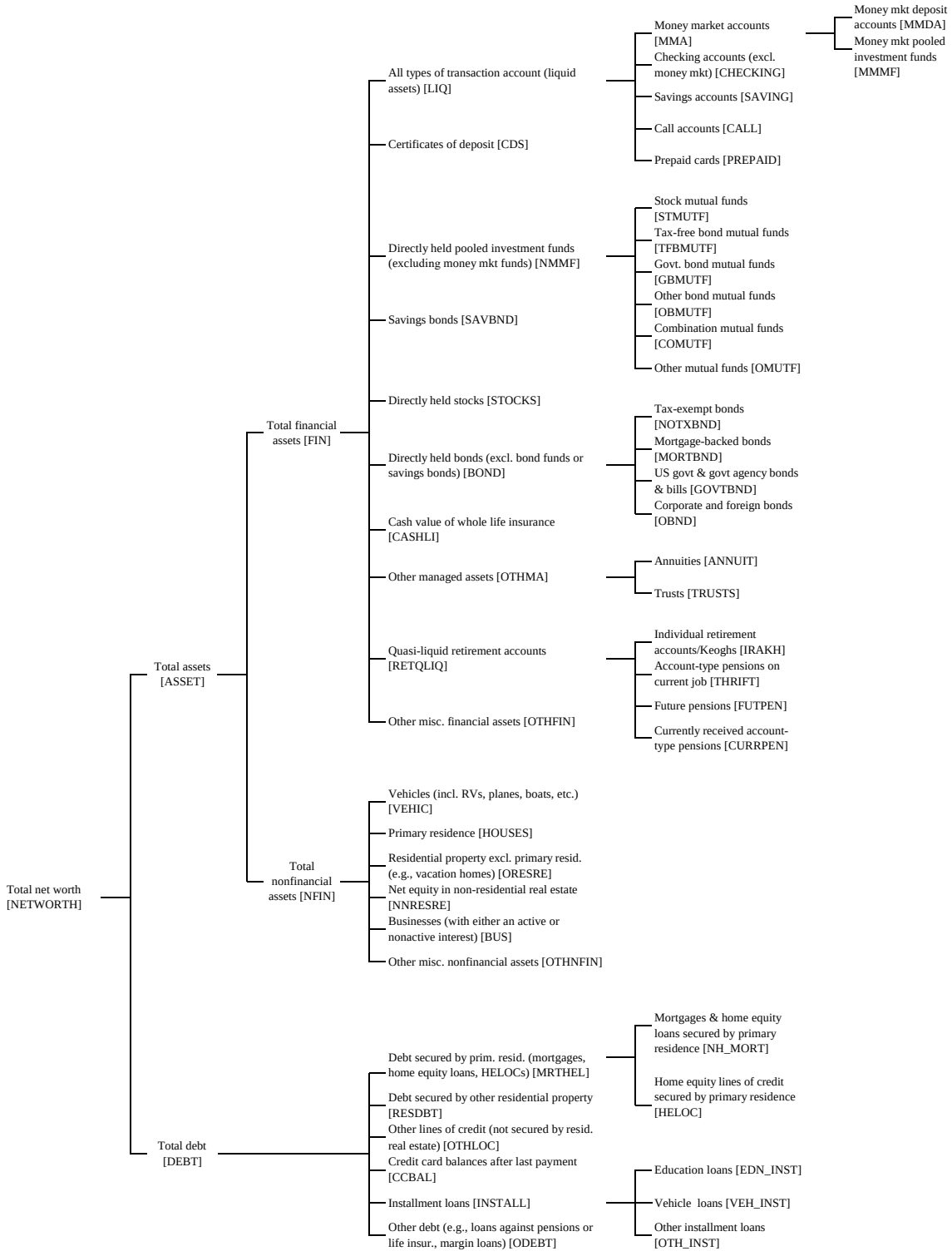

FIGURE A.1. Definition of wealth in the SCF. Names in brackets refer to variables in the SCF Bulletin extract data.

Source: https://www.federalreserve.gov/econres/scfindex.htm, accessed on October 13, 2019.

Appendix B. Sampling error of PUF estimates

In order to set the scene for the estimation of the PUF standard errors, I will first describe the key features of the INSOLE and PUF sampling designs. The reason for characterizing both designs is that the PUF is sub-sampled from the INSOLE, as opposed to being drawn directly from the underlying population of tax returns. Therefore, in order to fully characterize the PUF sampling design, it is necessary to start with a brief description of the INSOLE sample.

B.1. **INSOLE sampling.** The INSOLE is a disproportionate (highly) stratified probability sample of individual income tax returns. The returns are stratified using three variables: gross positive or gross negative income, presence or absence of special forms and schedules, and the return's potential usefulness for tax policy modeling (referred to as a return's *degree of interest*). Within each stratum, returns are selected for the sample using two sampling techniques, both of which are based upon the primary filer's Social Security Number (SSN).[1,2]

The first sampling technique focuses on the last four digits of the SSN. In this method, a return is selected for the INSOLE if and only if the last four digits of the primary filer's SSN match one of ten four-digit numbers chosen from the Social Security Administration's Continuous Work History Sample (CWHS).[3] This method, which I refer to as the CWHS selection, gives all taxpayers approximately a 1 in 1,000 chance of being selected, regardless of the strata they are assigned to.

The second sampling technique relies upon the "SSN transform," which is a uniformly distributed five-digit random number generated from the taxpayer's SSN. In this method, a return is selected for the INSOLE if and only if the taxpayer's SSN transform is less than or equal to the stratum-specific "sample number" given by

$$\kappa_j = 100,000 \times \left(s_j - 0.001 + 0.001 \times s_j\right) - 1, \tag{B.1}$$

where s_j denotes the prescribed sampling rate in stratum j.[4]

As indicated in Equation B.1, the sample number chosen to yield a $100s_j$ percent sample is set with an allowance for the CWHS selection. For illustration, consider

[1] Note that joint returns and married filing separately (MFS) returns have both a primary and secondary SSN, whereas all other types of returns have only a primary SSN.

[2] The SSN is a nine-digit national identification number issued to all US citizens as well as permanent and temporarily residents by the Social Security Administration (SSA). The SSNs issued prior to 2011 consist of three parts: the Area Number, the Group Number, and the Serial Number. The Area Number (the first three digits of the SSN) is determined by the zip code of the mailing address shown on the application for the SSN; the Group Number (digits four and five) is assigned based on the SSN's issuance date; lastly, the Serial Number (the last four digits) is chosen at random from the set of integers ranging from 1 to 9,999. Following a reform of the SSN Numbering Scheme from June 2011, all nine digits of the SSN are assigned randomly.

[3] Prior to 2005 there were only five CWHS endings in the INSOLE.

[4] The "-1" is a correction term that accounts for the fact that the range of the sample number is between 0 and 99,999. Otherwise, a sampling rate of 100 percent would yield a sample number of 100,000, which is outside the range.

stratum $j^\star$ with the prescribed sampling rate of $s_j^\star = 0.10$. It follows that returns in $j^\star$ have a CWHS selection probability of 0.001 and a probability of being selected based on the SSN transform being equal to 0.0991 ($= 0.10 - 0.001 + 0.10 \times 0.001$). This results in the sample number $\kappa_{j\star}$ equal to $9,909$ ($= 100,000 \times 0.0991 - 1$). Consequently, all returns from stratum $j^\star$ with an SSN transform less than or equal to 9,909 are selected into the sample.

Note that one of the most important features of the SSN transform is that it is constant across different tax years for a given taxpayer. This implies that if once selected for the INSOLE, the taxpayer continues to be selected for as long as he or she remains a primary filer and qualifies for a stratum with the same or higher sampling rate. On the contrary, if a taxpayer drops from a stratum with a 10 percent selection probability to one with a 5 percent selection probability, the probability of him or her being retained in the sample is equal to 50 percent.[5]

B.2. **PUF sampling.** The PUF is a disproportionate (highly) stratified probability sample of individual income tax returns. It is obtained by sub-sampling the INSOLE, which in turn is drawn directly from the underlying population of tax returns (see Section B.1). The INSOLE is sub-sampled for the PUF at different rates depending on two factors: the stratum from which a return was initially selected and the method used in the sampling process (the CWHS selection versus selection based on the SSN transform). Accordingly, the sub-sampling rates vary considerably within and across strata and range from zero (exclusion from the PUF) to one (drawn with certainty).

Across all strata, returns that are sampled for the INSOLE using the CWHS selection (as opposed to selection based on the SSN transform) are sub-sampled at a rate of 30 percent (or 70 percent following the PUF redesign from 2009).[6] In the first step of the two-step sub-sampling process, three (seven) numbers are drawn at random from a set of four-digit integers designated for the CWHS selection of the INSOLE. An important feature of this design is that once drawn at random, the three (seven) CWHS endings are retained in the sample. This implies that the same three (seven) four-digit integers were considered for the CWHS sub-sampling process every year between 1991 and 2008 (2009 and 2012). In the second step, a return is selected for the PUF if and only if the last four digits of the primary filer's SSN match one of the three (seven) numbers drawn in step one. This sub-sampling process gives all taxpayers a 3 in 9,999 (7 in 9,999) chance of being selected into the PUF, irrespective of the stratum. This compares to a 1 in 1,000 chance of being sampled for the INSOLE based on the CWHS selection process.

Returns that are sampled for the INSOLE using selection based on the SSN transform are sub-sampled for the PUF at different rates across different strata. The sub-sampling rates vary from 0 to 1 and can be classified into one of three main categories: drawn with certainty, sub-sampled at the rate that yields a 10 percent PUF

[5]More information on the INSOLE sampling design can be found in Czajka, Kirwan and Sukasih (2014).

[6]For more details regarding the redesign see Bryant, Czajka, Ivsin and Nunns (2014).

sampling rate, and excluded from the PUF.[7] For example, in year 2008, strata with gross positive or gross negative income between $250,000 and $1 million (in 1991 US dollars) were sampled with certainty; those with gross positive or gross negative income between $1 and $5 million were sub-sampled at rates that imply a 1 in 10 PUF sampling rate; and, returns with extreme values from strata with gross positive or gross negative income of over $1 million were excluded from the sample.

B.3. **Stratification of the INSOLE sample.** In the following, I present details about the stratification of the INSOLE sample. Specifically, in Section B.3.1, I discuss stratification by income, in Section B.3.2, stratification by the presence or absence of special forms and schedules, and in Section B.3.3 stratification by a return's degree of interest. Section B.3.4 characterizes two INSOLE priority strata, and Section B.3.5 concludes.

B.3.1. *Stratification by income.* As detailed in Czajka et al. (2014), the income measure used for stratification, say φ, is defined as the maximum between taxpayer's gross positive income and the absolute value of a taxpayer's gross negative income. Gross positive income is calculated as the sum of (i) twelve strictly positive items from Form 1040 and Schedule E and (ii) eleven items from Form 1040 and Schedules C, D, and F included only if the number is positive (see left panel of Table B.1). Similarly, gross negative income is calculated as the sum of (i) seven losses reported on Forms 1040 and 3903 and Schedules C and E, (ii) eleven items from Form 1040 and Schedules C, D, and F included (in absolute value) only if the number is negative, (iii) two deduction items from Schedules C and F, and (iv) three negative income adjustment items from Schedules C, E, and F (see right panel of Table B.1). Based on the constructed income measure φ, tax returns are stratified into one of nineteen income categories ranging from negative $10,000,000 or less to positive $10,000,000 or more, where the income classes are deflated using the Chain-Type Price Index for the Gross Domestic Product as of 1991.

The nine negative income classes are: $10 million or more (income level 1); $5–$10 million (income level 2); $2–$5 million (income level 3); $1 –$2 million (income level 4); $0.5–$1 million (income level 5); $250,000 –$500,000 (income level 6); $120,000–$250,000 (income level 7); $60,000–$120,000 (income level 8); under $60,000 (income level 9). The ten positive income classes are: under $30,000 (income level 10); $30,000–$60,000 (income level 11); $60,000–$120,000 (income level 12); $120,000–$250,000 (income level 13); $250,000–$500,000 (income level 14); $0.5–$1 million (income level 15); $1–$2 million (income level 16); $2–$5 million (income level 17); $5–$10 million (income level 18); and $10 million or more (income level 19).

B.3.2. *Stratification by degree of interest.* In addition to stratification by income, tax returns are further stratified by degree of interest, where "within the same income class, returns are considered more useful (or 'interesting') if less common income

[7] Prior to 2005, any stratum with an INSOLE sampling rate in excess of 1 in 3 was subsampled for the PUF at a rate of 1 in 3. Thus, if a stratum had an INSOLE rate of 35 percent, the PUF sampling rate would be a little over 1 in 9.

TABLE B.1. Items contributing to gross positive and negative income

Gross Positive Income	
Item	Source
Strictly Positive Items	
1. Wage amount	1040
2. Tax exempt interest	1040
3. Taxable dividends	1040
4. Alimony received	1040
5. Pension amount	1040
6. IRA distribution	1040
7. Unemployment compensation	1040
8. Social Security	1040
Strictly Gain Items	
1. Total rental payments amount	E
2. Total royalty payments amount	E
3. Partnership, s corporation income	E
4. Estate and trust income	E
Business Income Items (if positive)	
1. Schedule C-1 gross profit/loss	C
2. Schedule C-2 gross profit/loss	C
3. Schedule C-3 gross profit/loss	C
4. Schedule F-1 gross profit/loss	F
5. Schedule F-2 gross profit/loss	F
Net Items	
1. Supplemental gains/losses	1040
2. Other income amount	1040
3. Farm/rent income/loss	1040
4. Taxable interest income	1040
5. Net short-term gain/loss amount	D
6. Net long-term gain/loss amount	D

Gross Negative Income	
Item	Source
Loss Items	
1. Partnership, s corporation loss	E
2. Estate and trust loss	E
3. Total expenses all property amount	E
4. Total depreciation all property amount	E
4. Alimony paid	1040
6. Form 3903 moving expense amount	3903
7. Business-at-home expenses	C
Business Loss Items (if negative)	
1. Schedule C-1 gross profit/loss	C
2. Schedule C-2 gross profit/loss	C
3. Schedule C-3 gross profit/loss	C
4. Schedule F-1 gross profit/loss	F
5. Schedule F-2 gross profit/loss	F
Net Items	
1. Supplemental gains/losses	1040
2. Other income amount	1040
3. Farm/rent income/loss	1040
4. Taxable interest income	1040
5. Net short-term gain/loss amount	D
6. Net long-term gain/loss amount	D
Deduction Items	
1. Total deductions	C
2. Total farm expenses	F
Adjustment Items	
1. Negative income adjustment	C
2. Negative income adjustment	E
3. Negative income adjustment	F

Note: The letters C, D, E, and F denote a source tax schedule, and the numbers 1040 and 3903 a source tax form. Form 1040 is the standard IRS form used for filing purposes whereas Form 3903 summarizes moving expenses. Schedule C summarizes profits and losses from businesses, Schedule D capital gains and losses, Schedule E supplemental income and losses, and Schedule F profits and losses from farming. Source: Czajka et al. (2014).

sources or deductions are prominent" (Czajka et al., 2014). The degree of interest ranges from 1 (the least "interesting") to 4 (the most "interesting") and sub-stratifies four income classes with a gross positive income less than $\$250,000$. The income class with less than \$30,000 is sub-stratified into three classes whereas the other three income categories (i.e., $\$30,000 - \$60,000$, $\$60,000 - \$120,000$, $\$120,000 - \$250,000$) are sub-stratified into two. This yields a total of 24 distinct levels of income by degree of interest.

B.3.3. *Stratification by special forms.* The third and final dimension of stratification depends on the presence or absence of the following forms and schedules: Form 2555 (*Foreign Earned Income*), Form 1116 (*Foreign Tax Credit*), Schedule C, and Schedule F. In most years that are commonly referred to as *non-foreign study years*,

returns are classified into one of four mutually exclusive and collectively exhaustive categories based on a form type. The first category consists of returns with either Form 2555, Form 1116, or both; the second category consists of returns without Forms 2555 and 1116, but with Schedule C; the third category consists of returns without Form 2555, Form 1116, or Schedule C, but with Schedule F; finally, the fourth category consists of all other returns. In the remaining years (i.e., those ending in either 1 or 6, and commonly referred to as *foreign study years*), the first category is subdivided into two, one with returns with Form 2555 and the other with returns with Form 1116 but without Form 2555. This distinction between Forms 2555 and 1116 has the objective of over-sampling returns with Form 2555 in order to facilitate the use of the INSOLE in foreign income studies.

B.3.4. *Priority strata.* In addition to regular strata (i.e., strata generated by the combination of income level, degree of interest, and form type), there are two *priority* strata that take precedence over all regular strata and are sampled with certainty. This implies that a return is to be sampled from a regular stratum if and only if it does not qualify for any of the priority strata.

The first priority stratum (stratum 101) comprises returns with Adjusted Gross Income (AGI) or expanded income of at least $200,000 (in current dollars) and with no income tax liability after subtracting all credits. Therefore, this stratum is often referred to as one with "high-income non-taxable" returns.[8] Since the $200,000 income threshold is not deflated (using the Chain-Type Price Index for the 1991 Gross Domestic Product) but instead expressed in current dollars, the number of returns in this priority stratum has increased over time from 2,757 in 1991 to 4,114 in 2000 to 35,067 in 2012. The second priority stratum (stratum 201) comprises returns with business or profession receipts exceeding $50,000 (in current dollars). Even though the number of returns in this stratum increased over time (from 46 in 1991 through 1,025 in 2000 to 323 in 2012), it continues to be negligible, especially when compared to the number of returns in the other priority stratum.

[8]Expanded income is defined as AGI with "tax-exempt interest, nontaxable Social Security benefits, the foreign-earned income exclusion, and tax preference items used to calculate the alternative minimum tax", less "unreimbursed employee business expenses, moving expenses, investment interest expenses up to the value for investment income, and miscellaneous itemized deductions below the 2 percent of AGI floor" (Czajka et al., 2014).

TABLE B.2. Number of tax returns in the population, INSOLE, and PUF samples by stratum for 2008

Gross income	Degree of interest	Form 1040, with Form 1116 or Form 2555			Form 1040, with Schedule C but without Form 1116 or Form 2555			Form 1040, with Schedule F but without Schedule C, Form 1116 or Form 2555			Other forms			Total		
		Population	INSOLE	PUF	Population	INSOLE	PUF	Population	INSOLE	PUF	Population	INSOLE	PUF	Population	INSOLE	PUF
Negative income																
>=$10M	All	456	456	46	1,029	1,029	123	144	144	14	1,182	1,182	117	2,811	2,811	300
$5–$10M	All	860	860	78	1,724	1,724	175	260	260	16	2,202	2,202	231	5,046	5,046	500
$2–$5M	All	3,590	1,182	350	6,543	2,172	670	959	336	101	7,834	2,699	785	18,926	6,389	1,906
$1–$2M	All	7,462	1,203	756	13,451	2,126	1,318	2,358	388	241	15,571	2,436	1,492	38,842	6,153	3,807
$0.5–$1M	All	16,922	552	543	32,988	1,138	1,120	5,730	200	194	36,138	1,145	1,128	91,778	3,035	2,985
$250–$500K	All	33,620	350	325	74,321	715	658	11,769	115	106	80,568	806	757	200,278	1,986	1,846
$120–$250K	All	60,376	287	239	155,147	788	680	19,156	120	102	172,644	886	759	407,323	2,081	1,780
$60–$120K	All	70,010	196	154	201,319	649	479	20,470	72	54	249,179	785	589	540,978	1,702	1,276
<=$60K	All	61,200	108	70	473,691	888	535	28,417	51	30	800,133	1,469	884	1,363,441	2,516	1,519
Positive income																
<=$30K	**1**	⋆⋆	⋆⋆	⋆⋆	⋆⋆	⋆⋆	⋆⋆	⋆⋆	⋆⋆	⋆⋆	31,485,356	31,416	9,378	31,485,356	31,416	9,378
<=$30K	2	274,562	270	82	3,046,298	2,977	951	83,688	77	20	28,296,966	28,208	8,445	31,701,514	31,532	9,498
<=$30K	3-4	212,182	315	186	4,805,497	7,413	4,047	106,177	192	95	6,355,132	9,747	5,283	11,478,988	17,667	9,611
$30–$60K	1-2	681,524	678	209	1,978,448	2,028	585	171,290	168	46	21,950,763	21,915	6,598	24,782,025	24,789	7,438
$30–$60K	3-4	528,015	846	477	3,758,970	5,841	3,250	244,505	401	215	6,194,112	10,003	5,570	10,725,602	17,091	9,512
$60–$120K	1-3	1,085,559	1,091	336	2,301,810	2,299	672	217,320	242	73	11,353,034	11,257	3,373	14,957,723	14,889	4,454
$60–$120K	**4**	650,212	966	524	2,483,674	3,822	2,047	174,911	225	126	2,812,052	4,279	2,305	6,120,849	9,292	5,002
$120–$250K	1-3	337,092	634	413	401,367	756	503	82,868	188	111	1,265,376	2,473	1597	2,086,703	4,051	2,624
$120–$250K	4	822,634	2,741	2172	1,337,403	4,564	3656	88,221	296	220	1,816,816	5,888	4,628	4,065,074	13,489	10,676
$250–$500K	All	511,639	3,662	3,316	463,861	3,366	3,019	73,527	517	471	592,030	4,199	3,788	1,641,057	11,744	10,594
$0.5–$1M	All	230,080	5,675	5,513	134,302	3,397	3,316	27,545	642	628	157,455	3,908	3,801	549,382	13,622	13,258
$1–$2M	All	93,022	11,189	9,174	35,861	4,417	3,584	7,029	829	667	45,529	5,638	4,610	181,441	22,073	18,035
$2–$5M	All	44,770	14,457	4,512	11,754	3,807	1,189	1,836	590	172	16,142	5,175	1,609	74,502	24,029	7,482
$5–$10M	All	11,812	11,812	1,215	2,336	2,336	251	299	299	24	3,174	3,174	322	17,621	17,621	1,812
>=$10M	All	7,917	7,917	793	1,088	1,088	102	140	140	11	1,487	1,487	144	10,632	10,632	1,050
Total		5,745,516	67,447	31,483	21,722,882	59,340	32,930	1,368,619	6,492	3,737	113,710,875	162,377	68,193	142,547,892	295,656	136,343
Priority stratum 101														32592	32592	3278
Priority stratm 201														382	382	30
Grand total														142,580,866	328,630	139,651

Note: Table summarizes information for 93 regular strata (top panel) and 2 priority strata (bottom panel). ⋆⋆ denotes either that the data were combined or deleted or that no returns in the population had the characteristic or the characteristic was so rare that it did not appear on any of the sampled returns. $1M denotes 1 million US dollars and $1K 1 thousand US dollars. Sources: Internal Revenue Service Research (2010, 2012).

B.3.5. *Conclusion.* The combination of 24 levels of income by degree of interest and 4 form types yields 96 regular strata (= 24×4). However, since among returns with gross positive income less than \$30,000 only those without any special form or schedule can be assigned a degree of interest equal to 1 (which indicates the least "interesting" return), effectively, the number of regular strata in non-foreign study years is equal to 93 (= $23 \times 4 + 1$) and in foreign study years to 116 (= $23 \times 5 + 1$). To conclude, the INSOLE is sampled from a total of 95 (= $93 + 2$) strata in non-foreign study years, and from a total of 118 (= $116 + 2$) strata in the years 1991, 1996, 2001, and 2006. For illustration, see Table B.2, where I present the population, INSOLE, and PUF counts for each of the 95 strata for tax year 2008.

B.4. **PUF sampling error.** In order to estimate the PUF sampling error I first generate $L = 999$ bootstrapped sample replicates based on the publicly available information on taxpayers' strata and stratum-specific probability of selection. For brevity, let $\mathcal{S}$ denote the PUF sample of taxpayers, and assume that $\mathcal{S}$ comprises J mutually exclusive and collectively exhaustive strata such that

$$\mathcal{S} = \cup_{j=1}^{J} \mathcal{S}_j, \tag{B.2}$$

where $\mathcal{S}_j \cap \mathcal{S}_{j'} = \emptyset$ for all $j \neq j'$.

Moreover, let n denote the total sample size, and let n_j be the number of taxpayers selected for the sample from stratum j. Since the strata are mutually exclusive and collectively exhaustive, it follows from Equation B.2 that

$$n = \sum_{j=1}^{J} n_j. \tag{B.3}$$

Since across all tax years under consideration there exist strata with as low as 10 observations or fewer, bootstrapping methods cannot be applied directly to $\left\{\mathcal{S}_j\right\}_{j=1}^{J}$. Instead, I first classify the J strata into $J^\star \ll J$ clusters using the Partitioning Around Medoids (PAM) clustering procedure (see Reynolds, Richards, de la Iglesia and Rayward-Smith, 1992), where I determine the number of clusters in each tax year based on a silhouette analysis. The clustering procedure uses as an input three stratification variables originally designated for the INSOLE sample: gross income, presence or absence of special forms and schedules, and the return's potential usefulness for tax policy modeling. Since the income variable is ordinal (successive income brackets) whereas the latter two are nominal, I use the Gower distance measure, which is applicable to a mix of ordinal and nominal variables.

The clustering procedure results in a PUF sample of taxpayers $\mathcal{S}$ that comprises J^* mutually exclusive and collectively exhaustive clusters such that

$$\mathcal{S} = \cup_{j=1}^{J^\star} \mathcal{S}_j^\star, \tag{B.4}$$

where $\mathcal{S}_j^\star \cap \mathcal{S}_{j'}^\star = \emptyset$ for all $j \neq j'$.

Moreover, with $n_j^\star$ denoting the number of taxpayers in cluster j, it follows from Equation B.4 that

$$n = \sum_{j=1}^{J^\star} n_j^\star. \tag{B.5}$$

For example, in tax year 2008, I classify the 95 strata (with the minimum number of observations per stratum equal to 11) into 23 clusters (with the minimum number of observations per stratum equal to 184).

After classifying taxpayers into $J^\star$ clusters, I draw $L = 999$ independent bootstrapped sample replicates. Specifically, for each sample replicate $l : 1 \to L$, I draw with replacement $n_j^\star$ sample observations from each cluster j, such that the total number of observations in each sample replicate is equal to n.

Finally, let $\hat{\theta}$ denote the estimate of θ computed in the main data set, and let $\hat{\theta}_l$ denote the estimate of θ obtained in the lth bootstrapped sample replicate (as opposed to the main data set). I estimate the sampling error of $\hat{\theta}$ by a sample standard deviation of $\left\{\hat{\theta}_l\right\}_{l=1}^{L}$ as

$$\hat{\sigma}_{1,\hat{\theta}} = \sqrt{\frac{1}{L-1}\sum_{l=1}^{L}\left(\hat{\theta}_l - \frac{1}{L}\sum_{l'=1}^{L}\hat{\theta}_{l'}\right)^2}. \tag{B.6}$$

References

Bell, William R. and David W. Wilcox, "The effect of sampling error on the time series behavior of consumption data," *Journal of Econometrics*, 1993, *55*, 235–265.

Biemer, Paul and Lars Lyberg, *Introduction to survey quality*, Hoboken, New Jersey: John Wiley & Sons, 2003.

Bryant, Victoria L., John L. Czajka, Georgia Ivsin, and Jim Nunns, "Design changes to the SOI Public Use File (PUF)," *Prepared for the"New Resources for Microdata-Based Tax Analysis" Session of the 107th Annual Conference on Taxation, National Tax Association Santa Fe, New Mexico, November 15*, 2014, pp. 1–19.

Czajka, John L., Brendan Kirwan, and Amang Sukasih, "An assessment of the need for a redesign of the Statistics of Income Individual Tax Sample," *Mathematica Policy Research*, 2014, pp. 1–95.

Groves, Robert M., Floyd F. Fowler, Mick P. Couper, James M. Lepkowski, Eleanor Singer, and Roger Tourangeau, *Survey methodology*, Hoboken, New Jersey: John Wiley & Sons, 2009.

Internal Revenue Service Research, Statistics of Income, "Individual Income Tax Returns 2008," *Publication 1304 (Rev. 07-2010)*, 2010, pp. 1–295.

_____, "General description booklet for the 2008 Public Use Tax File," 2012, pp. 1–42.

Kennickell, Arthur B., "Revisions of the variance estimation procedure for the SCF," *Working Paper of the Board of Governors of the Federal Reseve System*, 2000.

_____, "The role of over-sampling of the wealthy in the Survey of Consumer Finances," *Irving Fisher Committee Bulletin*, 2008, *28*.

Meng, Xiao-Li, "Multiple-Imputation inferences with uncongenial sources of input," *Statistical Science*, 1994, *9* (4), 538–558.

Reynolds, Alan P., Graeme Richards, Beatriz de la Iglesia, and Victor J. Rayward-Smith, "Clustering rules: A comparison of partitioning and hierarchical clustering algorithms," *Journal of Mathematical Modelling and Algorithms*, 1992, *5*, 475–504.

Tourangeau, Roger, Robert A. Johnson, Jiahe Qian, Hee-Choon Shin, and Martin R. Frankel, "Selection of NORC's 1990 National Sample," *Working Paper, National Opinion Research Center at the University of Chicago*, 1993.

Yang, Shu and Jae K. Kim, "A note on multiple imputation for method of moments estimation," *Biometrika*, 2016, *103* (1), 244–251.